Briefing to NOAA leadership on recommendations for statistical post-processing

Tom Hamill (ESRL/PSD); Barb Brown (NCAR), Kathy Gilbert (WPC/OPC),
Matthew Peroutka, Dave Rudack (MDL), Yuejian Zhu (EMC)

With thanks to ~90 other participants from universities, foreign weather services, and private companies.

January 2016

You are here: NCAR • RAL • Events • *The Future of Statistical Post-processing in NOAA and the Weather Enterprise*

The Future Of Statistical Post-processing In NOAA And The Weather Enterprise

About | Agenda | Logistics | Registration | Contacts

THE FUTURE OF STATISTICAL POST-PROCESSING IN NOAA AND THE WEATHER ENTERPRISE

January 19 - 22, 2016

NOAA Center for Weather and Climate Prediction Building
5830 University Research Ct, College Park, MD 20740

The NOAA / National Weather Service's (NWS) Strategic Plan, "Weather-Ready Nation" envisions NOAA providing dramatically improved weather decision support to its customers through the generation of much more skillful and reliable forecast guidance. Statistical post-processing of weather forecast guidance (its adjustment using discrepancies between past forecasts and observations/analyses) will be a key component of helping the NWS meet this goal.

Statistical post-processing is aided tremendously when the relevant data is in place, including recent or retrospective forecasts (reforecasts) and associated recent or retrospective observations and analyses. NOAA intends to generate these data sets more regularly and make them of higher quality and consistency in the future.

This workshop will help NOAA set its future requirements for providing internal and external customers with the high-quality data they need to achieve expected benefits from statistical post-processing. The workshop will discuss several topics, including: (1) a conversation with internal and external customers about their desires and requirements for data. For example, what length of retrospective forecast and analysis data is needed? What storage formats are most convenient? What regions of the world are needed, and what weather elements are most important to save? What platforms should host the data, internal NOAA, and/or commercial cloud vendors' platforms? (2) Conversations about how the post-processing community can make more rapid progress through sharing of software, including potentially establishing a community repository of post-processing algorithms, test data sets, I/O routines, verification routines, and so forth. (3) New developments in statistical post-processing, especially ones that may have implications with regards to what data NOAA should store and make available.

The Future of Statistical Post-processing in NOAA and the Weather Enterprise

Held:

January 19 - 22, 2016

NOAA Center for Weather and Climate Prediction Building
5830 University Research Ct
College Park, MD 20740

Workshop Sponsors

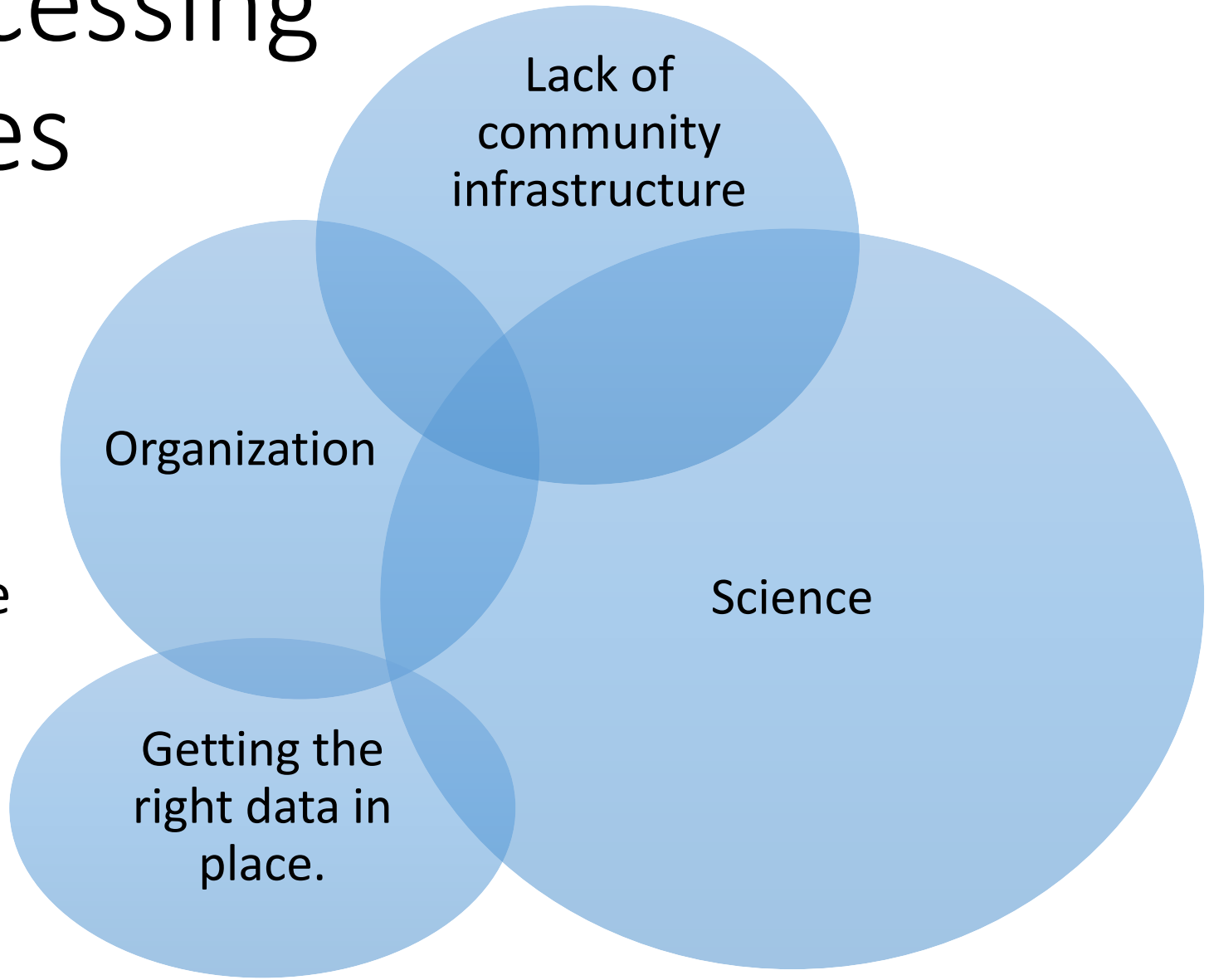
For more information, please contact Tom Hamill (tom.hamill@noaa.gov) or Dave Rudack (David.Rudack@noaa.gov).

This briefing synthesizes the recommendations from a recent workshop on statistical post-processing, held in College Park MD, Jan 2016.

This meeting helped inform NGGPS post-processing team plans, but also was meant to facilitate planning extramural to NGGPS.

Post-processing challenges

Note the substantial overlap; e.g., some science challenges are also organizational challenges.



Science challenges

- The methodology behind our current products isn't consistently statistically rigorous.
  - Atmospheric scientists are mostly performing the product development.
  - Better products are possible with input of professional statisticians.
  - Preferred methodology may change as training data is improved and lengthened.
- NOAA has multiple overlapping products produced with a variety of methods, and we haven't carefully evaluated strengths/weaknesses.
  - An organizational challenge also.
- Need reforecasts of high quality and statistical consistency.
- Need observation / reanalysis training data of high quality.
  - If analyses are to be used as surrogate for truth, they must be unbiased and low in error, else product quality suffers.

Organizational challenges

- Parallel product development in several organizations.

- Existing post-processing infrastructure is complex and hard to maintain.
  - Loss of productivity as we engineer around 20-year old software.
  - Not set up for new era of reforecasts.

Data challenges (1)

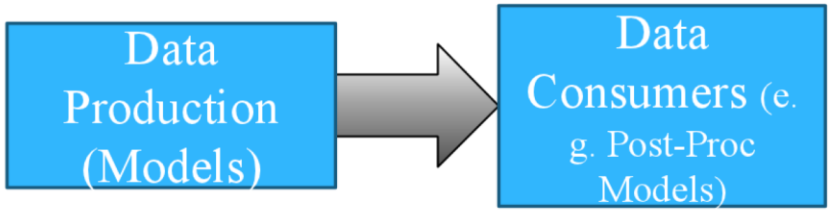
Room for the new reanalysis and reforecast kids ?

- Data storage: NOAA lacks a coherent plan to gather, store, and use the data sets required for statistical post-processing. Each group seems to assemble an archive that suits their own needs.
  - Want lots of forecast, observation, analysis data stored on disk, and we want it able to be read into post-processing algorithms quickly; private sector too.
  - Data storage demands of post-processing not reflected in requirements for future HPC purchases.
- Community agrees reforecasts needed (science challenge as well).
  - Quality reanalyses needed too.
  - Reanalysis/reforecast not reflected in requirements for future HPC purchases.
  - When has system changed enough to need a new reforecast? Science challenge also.
- A data storage plan that's convenient for NOAA internal development may not be convenient for industry, academia, inter-governmental collaboration.

Data challenges (2)

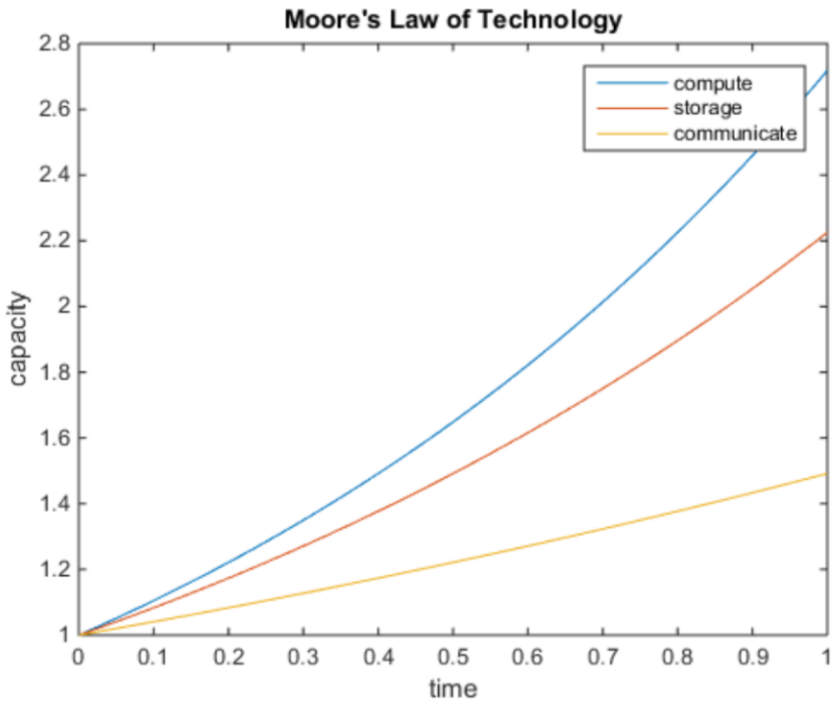
- Performing the product development on different systems from the production of data will cause greater and greater challenges over time, as CPU increases faster than disk and communication.

Our national data flow model



Bring the data to the end uses

Is not scalable and is fundamentally broken



slide c/o Peter Neilley, the Weather Company

Little sharing of post-processing software and test data.

(Mostly) unconnected silos of software development; hard to find other's code; sub-standard documentation; unclear policies on use of GitHub and other public repositories; lack of modern version control; training data hard to get.

More on the software development conundrum.

- Runs on the operational WCOSS system.
- Developed on WCOSS also to facilitate implementations.
- Internal data formats used like “TDLPack”
- Code base must be protected.

- None of ESRL/PSD nor external community development and testing is on WCOSS.
- Different data formats common, e.g., reforecasts packaged into netCDF files.

Recommendations for these topics:

- Improving NOAA's post-processing science.
- Organizational changes.
- Building a community infrastructure.
- Putting the data in place.

Potential cost savings highlighted in RED. Risks in PURPLE

Improving the science (1)

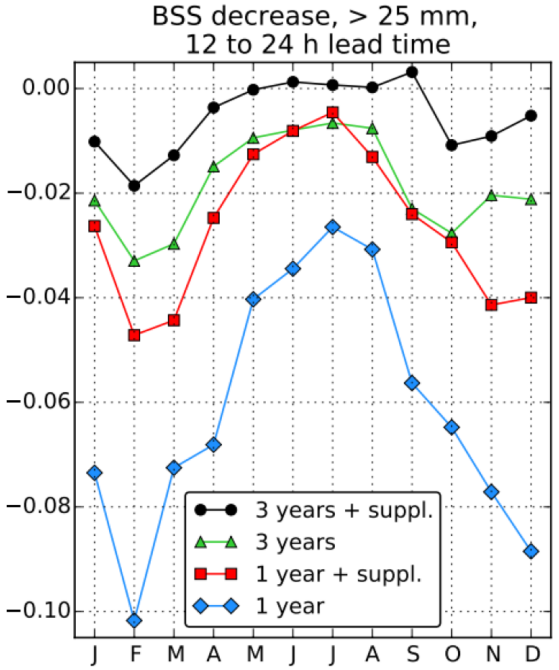
- Hire ~4 professional statisticians to work with atmospheric scientists.
  - UMAC recommended 4 Ph.D-level statisticians at MDL, with at least one ASAP.
  - Interim measure? Summer visitors from prominent university programs.
  - Existing staff will be more productive with oversight.

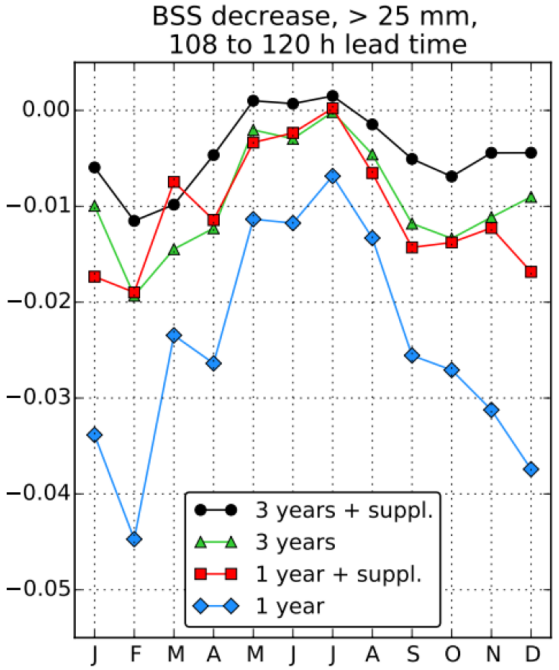
Improving the science (2)

- Perform systematic comparisons of existing algorithms, and use the best across applications and organizations to improve seamlessness.
  - Define the methods of evaluation.
  - Determine parameters (e.g., precipitation) where unified code possible.
  - Re-code the best algorithm(s) so that they can be applied broadly across prediction systems and time scales.
  - Test for CPC, MDL, WPC, EMC, etc. applications
  - Decide.
  - Deploy.
  - Improved software maintainability can lead to a more efficient use of resources throughout the organization.

Improving the science (3)

- Investigate methods that promise equal (or greater) skill with less training data using extensive reforecast data available for GEFS v10.





Example of more efficient use of short training data sets with “supplemental locations”

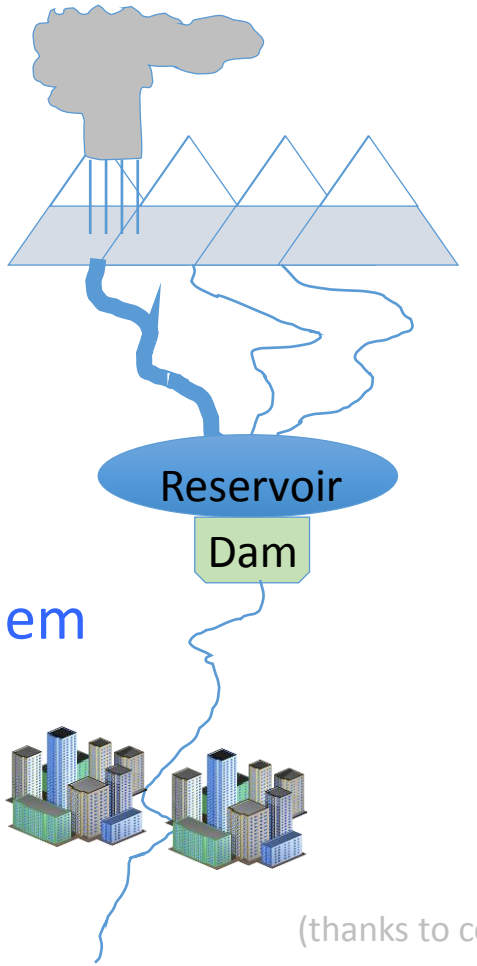
The baseline skill for precipitation forecasts is one trained with 11 years of data (for CONUS, 1/8-degree resolution). The degradation in skill with lesser training data is shown here.

3 years of data and supplemental locations produces skills that are only marginally worse than 11 years (without supplemental locations).

After we determine which post-processing method is the best, more systematic exploration along these lines will help us understand how to minimize reforecast computational expense in the future.

from Michael Scheuerer presentation at the post-processing workshop.

Improving the science (3) : Investment in the development and evaluation of post-processing methods for multi-variate calibration.



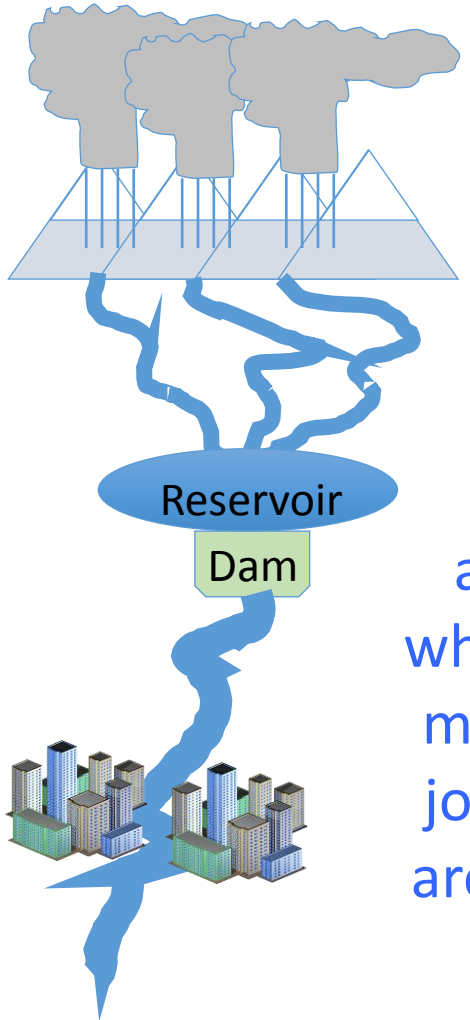
No problem

Example: hydrologists want to know not only the intensity of rainfall, but whether or not that intense rainfall will fall simultaneously in many nearby sub-basins.

What is the "copula" structure, i.e., the joint probabilities?

Decision makers save with better decisions.

(thanks to conversations with John Schaake)



a problem when either marginal or joint probs. are not well forecast

Other major science recommendations

- Proceed briskly to probabilistic NDFD, with improved customer decisions.
  - Risk: more data transmission and local storage cost.
- Improve data assimilation methods for producing high-quality, high-resolution, unbiased analyses for training and verification.
  - Improved products → improved customer decisions.
- Determine potential value of future reforecasts for regional models using short-range forecasts from global models.
  - Also, compare to post-processing products based on short training regional model datasets (e.g. MOS)

Organizational recommendations: (1)

- Improve the procedures for defining statistical post-processing requirements, following UMAC recommendations.
- Develop strategic plan and roadmap for post-processing across NOAA over 5-year time scale. Use it.
  - Risk: many such plans in NOAA in past haven't been executed.
- Again, employ ~4 professional statisticians / consultants to raise quality of science.
  - in the interim, some university collaborators / visitors?
  - makes existing scientists more productive.

Reorganization possibilities. (not mutually exclusive, and ordered major to minor)

- MDL becomes the NOAA center for production of the base post-processing elements. Move some WPC, CPC, HPC, EMC products to MDL (with augmented / reallocated MDL staff). Eventual cost savings from reducing duplication.
  - MDL would need greater scientific expertise to become a true center of excellence.
  - WPC post-processing concentrates on techniques for incorporation of human guidance.
  - EMC focuses on production of forecasts, analyses.
  - Risk: MDL has its hands full.
- New post-processing testbed for accelerating R2O, testing advanced techniques.
  - Risk: have other test beds produced expected return on investment?
- Continue distributed post-processing development, but with some centralized oversight, i.e., a post-processing "czar."
  - Risk: what if czar and organizational managers differ?
- Train staff in software best practices, modern languages; set management expectation across NOAA for collaboration, use of modern community software.
  - http://software-carpentry.org/
  - Much more rapid rate of software improvement.

Organizational recommendations (3)

- MDL to re-engineer post-processing support infrastructure, such as the archival of training data (see data recommendations later). "WISPS"
  - Create requirements and guidelines for meta-data for forecasts, reforecasts, and post-processed data.
  - Reduced software maintenance cost.
- Develop a community support infrastructure (also community infrastructure): Options:
  - MDL to build and maintain a community hub for software, test data, verification.
  - A new NOAA test bed for post-processing with this responsibility.
  - Leverage developments elsewhere, more rapid product improvement.

Community infrastructure recommendations (1): Build a community repository, and …

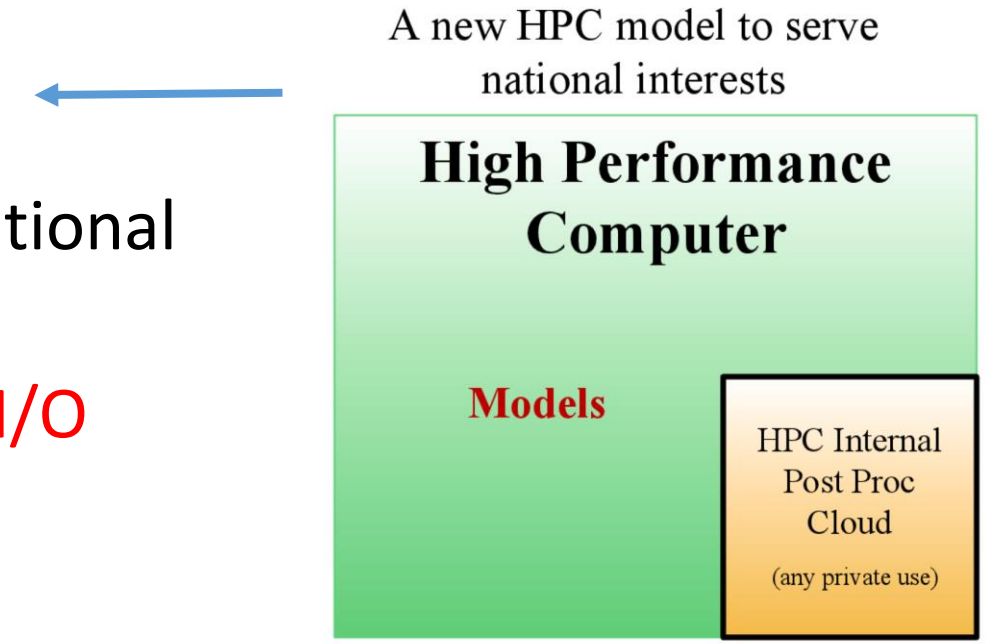
- Establish modern, widely-utilized distributed version control system to facilitate R2O and O2R.
  - Git and GitHub are emerging industry standards.
  - Potential tiers in the repository (inner: NOAA, closely maintained, to outer, with comparative ease for external collaborators to use and modify).
- Institute a ticket-tracking system to monitor requested product improvements and their disposition.
- Establish process and allocate resources to manage incorporation of external code contributions into the main repository.
- Establish requirements for metadata and documentation.
- Establish a centralized location for documentation and data access (e.g., data.gov).

Community infrastructure recommendations (2)

- Establish governance plans to enable groups to work together and methods of making decisions.
- Leverage work already done by 18F, as well as work with them on creating guidelines and standards.
- Work with the community to determine two or three modern common data formats that should be used (e.g., netCDF, HDF, geoJSON) that will satisfy operational, research, collaboration, and archival purposes.
- Provide assistance to collaborators.

Community infrastructure recommendations (3)

- Create an area where data can be accessed and processed by external collaborators easily, including international collaborators.
- Convenient workaround for storage, I/O bottleneck, especially for corporate partners.
- Risk: can security concerns be addressed?



proposal c/o Peter Neilley, the Weather Company.

Data and parameter requirements recommendations: (1)

- NOAA must post-process and make available “foundational data” for use in NOAA and across the broader enterprise.
  - Private companies heavily use NOAA post-processed guidance.
  - Some users note desire for broader, 3-D, global post-processed guidance.
  - Foundational data: variables like temp, precip, others in NDFD.

- Quality **global** data sets (reanalysis, reforecast, post-processed, satellite, radar, point data) should be made available given NOAA and private company global interests.
  - public servers and/or cloud.
  - Risks: computational expense, questionable quality in data-sparse regions.

Data and parameter requirements recommendations: (2)

- Follow up workshop with a structured survey of post-processing product developers to make sure we are saving on disk the relevant predictor information.
  - Some notes on preliminary survey are in the supplementary slides.

- Requirements for products should incorporate predictability considerations.
  - Example: production and storage of day +7 hourly precipitation forecasts at 1 km not warranted.

Links to specific breakout group facts, findings, and recommendations

- Methodology group here.
- Community infrastructure group here.
- Data/requirement group here.

- Folder with all post-processing workshop material here.